Problem 1.

(1.) Markov's Inequality is tight

Consider a random variable $X = \begin{cases} 1, & \text{with probability } \frac{1}{k} \\ 0, & \text{with probability } 1-\frac{1}{k} \end{cases}$, for any $k \geq 1$.

In this way, the mean $\mu = \frac{1}{k}$, and hence:

$$Pr[X \geq kE[X]] = Pr[X \geq 1] = \frac{1}{k}$$

(2.) Chebyshev's Inequality is tight

Let X be the random variable with the values $X = \begin{cases} -1, & \text{with probability } \frac{1}{2k^2} \\ 0, & \text{with probability } 1-\frac{1}{k^2} \\ 1, & \text{with probability } \frac{1}{2k^2} \end{cases}$, for any $k \geq 1$.

For this distribution, the mean $\mu = 0$ and the standard deviation $\sigma = \frac{1}{k}$, so

$$Pr(|X-\mu| \geq k\sigma) = Pr(|X| \geq 1) = \frac{1}{k^2}.$$

For any $k \geq 0$, this distribution is tight for **Chebyshev's inequality**.

Problem 2.

If $I = J$, then the summation is $(-1)^0 = 1$, so the formula is true for the first case.
If $I \neq J$, let $k = |J| - |I|$, then the summation can be rewritten as:

$$(-1)^0 \times \binom{k}{0} + (-1)^1 \times \binom{k}{1} + (-1)^2 \times \binom{k}{2} + ... + (-1)^k \times \binom{k}{k} = \sum_{i=0}^{k} (-1)^i \binom{k}{i} = (1-1)^k = 0$$

where the last equality is from binomial theorem:

$$(1+x)^n = \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + ... + \binom{n}{n}x^n = \sum_{i=0}^{n} x^i \binom{n}{i}$$

Problem 3.

Let $F = \{S_1, S_2, ..., S_m\}$ be a k-uniform k-regular family. To apply LLL, we try to define bad events and then prove that the probability that no bad events happen is greater than 0. (i.e. $Pr[\overline{A_i}...] > 0$). Therefore, we let A_i denote the event that S_i is not 2-colorable and prove that it is possible that no S_i is not 2-colorable.

$Pr[A_i] = 2 \times 2^{-|S_i|} = 2 \times 2^{-k} = 2^{1-k} = p$, and there are at most k^2 other sets in F that could intersect with S_i, so the maximum degree in the adjacency graph is $d = k^2$.

We can then calculate $ep(d+1) = e \times 2^{1-k} \times (k^2+1)$.
For $k = 10, e \times 2^{-9} \times 101 < 1$, and since 2^k decreases faster, so $ep(d+1) \leq 1$ holds for all $k \geq 10$.

Problem 4.

Let $x(A_i) = \frac{1}{d+1}$, then from the assumption that $ep(d+1) \leq 1$, we have:

$$ep(d+1) \leq 1 \rightarrow p \leq \frac{x_i}{e} \leq x_i(1-\frac{1}{d+1})^d = x_i \Pi_{j:(i,j)\in E(G)}(1-x_j)$$

From **Theorem 19.2**, if $\exists\ 0 \leq x_i < 1\ \forall\ 1 \leq i \leq n$, such that $\Pr[A_i] \leq x_i \Pi_{j:(i,j) \in E(G)}(1 - x_j)$, then $\Pr[\bigcap \overline{A_i}] > 0$

Problem 5.

1. Show that $Y_i = 0$ and $Y_i = 1$ with probability $\frac{1}{2}$ each.

$\Pr[Y_i = 0] = \Pr[\text{Both 0 or Both 1}] = \frac{1}{2} \times \frac{1}{2} \times 2 = \frac{1}{2}$

$\Pr[Y_i = 1] = \Pr[\text{One 0 and One 1}] = \frac{1}{2} \times \frac{1}{2} \times 2 = \frac{1}{2}$

2. Show that $E[Y_iY_j] = E[Y_i]E[Y_j]$ for any $1 \leq i, j \leq m$ and derive Var[Y].

Let $X = Y_iY_j$, then $\Pr[X = 0] = \Pr[Y_i = 0 \vee Y_j = 0] = \frac{3}{4}$, and $\Pr\ [X = 1] = \Pr[Y_i = 1 \wedge Y_j = 1] = \frac{1}{4}$.

So we have: $E[X] = \frac{3}{4} \times 0 + \frac{1}{4} \times 1 = \frac{1}{4} = E[Y_i]E[Y_j]$

From Figure 1:
If $E[X_iX_j] = E[X_i]E[X_j]$ for every pair of i and j with $1 \leq i,\ j \leq n$, then $\mathrm{Var}[X] = \sum_{i=1}^{n} \mathrm{Var}[X_i]$.

Problem 2. *(Exercise 3.15 from MU)* Let the random variable X be representable as a sum of random variables $X = \sum_{i=1}^{n} X_i$. Show that, if $\mathbb{E}[X_iX_j] = \mathbb{E}[X_i]\mathbb{E}[X_j]$ for every pair of i and j with $1 \leq i < j \leq n$, then $\mathbf{Var}[X] = \sum_{i=1}^{n} \mathbf{Var}[X_i]$.

Solution: From the definition of variance, we write

$$\begin{aligned}
\mathrm{Var}[X] &= \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i - \mathbb{E}\left[\sum_{i=1}^{n} X_i\right]\right)^2\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}(X_i - \mathbb{E}[X_i])^2 + 2\sum_{i<j}(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])\right] \\
&= \sum_{i=1}^{n} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathbb{E}[X_iX_j - \mathbb{E}[X_i]X_j - X_i\mathbb{E}[X_j] + \mathbb{E}[X_i]\mathbb{E}[X_j]] \\
&= \sum_{i=1}^{n} \mathrm{Var}(X_i)
\end{aligned}$$

Since

$$\mathbb{E}[X_iX_j - \mathbb{E}[X_i]X_j - X_i\mathbb{E}[X_j] + \mathbb{E}[X_i]\mathbb{E}[X_j]] = 2\mathbb{E}[X_i]\mathbb{E}[X_j] - 2\mathbb{E}[X_i]\mathbb{E}[X_j] = 0$$

Figure 1: Variance of a random variable as a sum of random variables

So, $\mathrm{Var}[Y] = \sum_{i=1}^{m} \mathrm{Var}[Y_i] = \frac{n(n-1)}{8}$

3. Use Chebyshev's inequality to derive a bound on $\Pr[|Y - E[Y]| \geq n]$.

$\Pr[|Y - E[Y]| \geq n] \leq \frac{Var[Y]}{n^2} = \frac{(n-1)}{8n}$

Reference

[1]`https://www.cs.ox.ac.uk/people/varun.kanade/teaching/CS174-Fall2012/HW/HW3_sol.pdf`